DOCUMENT RESUME

ED 319 081 CS 507 155

AUTHOR Dudczak, Craig A.; Day, Donald L.
TITLE A Profile of CEDA Debate Critics.
PUB DATE Aug 89
NOTE 36p.; Paper presented at the Annual Conference on Argumentation (Alta, UT, August 1989). For the second part of this study, see ED 318 071.
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Communication Research; Correlation; *Debate; Debate Format; Evaluation Criteria; Higher Education; *Judges; Profiles; Questionnaires; Speech Communication
IDENTIFIERS Critics; Cross Examination Debate Association; *Debate Tournaments

ABSTRACT

A study was conducted to develop a taxonomy of decision criteria employed by Cross Examination Debate Association (CEDA) debate critics. Four hypothesis characteristics were tested: (1) audience-centered critics would have a higher proportion of presentational remarks than analytic-centered critics; (2) audience-centered critics would devote a higher proportion of their ballots to critique than analytic-centered counterparts; (3) elimination round ballots for all critics would contain a higher proportion of decision criteria to critique than preliminary round ballots; and (4) elimination round ballots would have a higher proportion of substantive elements than preliminary round ballots. Subjects were 13 debate critics who had judged intercollegiate debate for fewer than 6 years. The study integrated structured data (questionnaires and template portions of debate ballots) with unstructured data (written portions of debate ballots and judging philosophies). Results of correlational analysis found only two instances of professed preference from the questionnaire corresponding with actual preferences and indicated that several clusters of ballot behavior were indicated by different paradigms. Results showed that philosophy statements were better predictors of both ballot behavior and survey responses than questionnaire responses. Support was not found for hypotheses one and two. Support was obtained for hypotheses three and four, indicating that critics reduce the amount of critique devoted to their written ballots in elimination rounds compared with the amount devoted in preliminary rounds. (Eight tables of data are included. Appendixes include the judging criteria questionnaire, coding categories for ballot comments, and judge philosophy coding categories. Twelve references are attached.) (MG)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED319081

A PROFILE OF CEDA DEBATE CRITICS

Craig A. Dudczak
Speech Communication
Syracuse University

Donald L. Day
Information Studies
Syracuse University

Sixth SCA/AFA Conference on Argumentation
Alta, Utah
August 1989

CS507155

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

CRAIG A. DUDCZAK

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

A PROFILE OF CEDA DEBATE CRITICS

Craig A. Dudczak
and
Donald L. Day
Syracuse University

The study in progress is an attempt to develop a taxonomy of decision criteria employed by CEDA debate critics. The study is exploratory in at least two dimensions. First, the study attempts to associate professed judging philosophy and responses to survey questions with ballot behavior. While judging philosophies and preferences expressed through survey instruments may be taken as "ought" statements, the pattern of decision criteria employed on ballots constitutes actual practice. One expects to find consistency between the philosophies and preferences judges express on the one hand and their comments to debaters and reasons for decision on the other.

Second, the study attempts to develop "judging profiles." Unlike NDT debate (characterized by fairly well-articulated "paradigms"), CEDA debate offers less well-defined (let alone accepted) perspectives regarding how rounds should be evaluated. The development of "judging profiles" is an attempt to discover (1) whether tacit paradigms exist and (2) what elements these paradigms contain. A taxonomy of debate critics would allow standardized review of judges' work products (ballots and philosophies) and would encourage development of sound principles of criticism on ballots. The taxonomy also would assist educators in organizing and conducting debate training.

This manuscript reports the first part of the study. The

analysis is limited to reporting the correspondence among preferences expressed through judge philosophy statements, responses to a survey instrument, and comments/decision criteria expressed on debate ballots. The emergent "judging profiles" will be reported in a subsequent manuscript.

The justification for this investigation may be found in the scarcity of information we possess about debate critic decision criteria. Previous researchers attempted to determine whether judging behavior corresponded with assumptions of decision paradigms. The earliest investigations (Cox 1974; Cross & Matlon 1978; Thomas 1977) were limited to NDT debate. They shared a limitation common to subsequent surveys (Buckley 1983; Lee, Lee & Seeger 1983; Gaske, Kugler & Theobald 1985) in that they relied exclusively on self-report. While data acquired by such means may reflect prevailing attitudes within the forensic community, they do not validate whether reported preferences actually are applied as criteria in the resolution of debate rounds. Moreover, the Gaske, Kugler, and Theobald research, while attempting to discriminate among CEDA judging paradigms, relied upon unequal (and generally subcritical) cell sizes which violate the assumptions of parametric statistics (61-65). Judges may have articulated perspectives in instruments used for any of these studies which they subsequently violated in their judging behavior.

Only two studies have taken the ballot artifacts of debate as the basis for analysis. Bryant (1983) compared selected NDT and CEDA debates to analyze the application of evidence within

CEDA and NDT formats. His results are contaminated, however, by a failure to control for differences in time format and for competitors' varying skill levels. Also, Bryant compared unequal debate experience levels (3-4).

Hollihan, Riley, and Austin (1983) investigated "themes" differentiating CEDA critics from their NDT counterparts. They employed content analysis to compare ballots written by judges in the two debate formats. Their results supported the existence of different "visions" embraced by CEDA judges vs. those in NDT.

Nevertheless, Hollihan et al limited comparisons between judges and their decision criteria in two important ways. First, they treated CEDA (and NDT) judges as monotheistic. Their analysis presumed CEDA judges were of one type. This assumption is suspect at least when applied to NDT judges because it is commonly held that competing paradigms are operating. There is also reason to expect that varying judging perspectives are applied in CEDA.

Second, Hollihan et al only looked at ballot comments as their artifact. Without knowledge of an individual judge's prior preferences regarding debate practices or theory, one cannot determine whether the absence of ballot comments reflects debater adaptation to the critic or inconsistency on the part of the judge. At the time of the Hollihan et al research, CEDA had not yet instituted a national tournament finals, with its judge philosophy booklet. NDT had employed judge philosophies since the 1970s. Differences in ballot comments reported by Hollihan et al may reflect, in part, greater availability of judging

preference statements for NDT judges.

Now that CEDA has institutionalized the practice of compiling judge philosophy statements, analysis can turn to investigating whether a correspondence exists between (1) what judges profess to employ as judging criteria and (2) their actual bases for decision, as reflected through ballot behavior.

Formal criticism offered during competition is a key feature of intercollegiate debate training. In particular, rationales offered in justification of decisions play a major role in shaping the aspects of debate emphasized subsequently by coaches and participants.

However, such criticism (offered as comments on ballots) is anecdotal. Many variables, some unrelated to standards of debate, may be applied. Since ballots are the primary feedback for debaters, an attempt should be made to describe criteria applied to decisions in a more systematic fashion.

The study in progress is guided by two overarching questions. Research question #1 asks "What is the strength of relationship between professed reasons for decision as claimed in a questionnaire and actual reasons cited in debate ballots?" This question broadly asks whether the elements disclosed as preferences through the survey instrument are used in resolving debates. Research question #2 asks "What is the strength of relationship between professed judging paradigms as claimed in a questionnaire and reasons for decision cited in debate ballots?" This question more narrowly asks whether debate judges who employ the same label display similar behaviors in the criteria they

employ to judge debates. One would expect debate judges who identify with a debate paradigm to employ decision criteria similar to those of others who claim to prefer the same paradigm.

Although the present study is intended only as a pilot, researchers opted to test four hypotheses characteristic of those that might be examined in a larger project.

Hypothesis #1--"The mean proportion of presentational (vs. substantive) remarks on ballots by Audience-centered critics (Argument Skills, Argument Critic, Public Audience) will be equal to the proportion of such remarks made by Analytic-centered (Value Comparison, Policy Implications, Stock Issues, Hypothesis Testing, & Judicial Model) critics." This hypothesis (stated to reject the null hypothesis for confirmation) expected that Audience-centered critics would have a higher proportion of presentational remarks than Analytic-centered critics.

Hypothesis #2--"The mean proportion of ballots devoted to critique (vs. decision criteria) by Audience-centered critics will be equal to the proportion allotted by Analytic-centered critics." This hypothesis expected Audience-centered critics to devote a higher proportion of their ballots to critique than Analytic-centered counterparts.

Hypothesis #3--"The mean proportion of ballots devoted to decision criteria (vs. critique) on elimination round ballots will be equal to the proportion allotted in preliminary rounds." This hypothesis expected that elimination round ballots for all critics (regardless of paradigm preference) would contain a higher proportion of decision criteria to critique than

preliminary round ballots.

Hypothesis #4--The mean proportion of substantive (vs. presentational) remarks made on ballots in elimination rounds will be equal to the proportion of such remarks made in preliminary rounds." This hypothesis expected that elimination round ballots would have a higher proportion of substantive elements than preliminary round ballots.

METHOD

The present study integrates structured data (from the questionnaire and from the template portions of debate ballots) with unstructured data (from written portions of ballots and judging philosophies). The advantage of using both survey research and content analysis is that the two techniques generate complementary findings which are more valid than those obtained when using either alone (Paisley 1969; Webb and Roberts 1969). Options offered in structured instruments reflect preconceptions held by the researcher. In other words, the respondents' choices are dictated by the instrument. Content analysis, on the other hand, begins with a view of reality held by the subject and attempts to conform that world to the analytic scheme of the researcher (Holsti 1969; Krippendorff 1980).

Subjects:

Subjects used in this pilot study were debate critics who judged debate rounds at CEDA tournaments primarily in the Northeast during the 1988-89 season. Many had judged intercollegiate debate for fewer than six years (63.1%), although most critique the equivalent of one tournament per semester

(94.7%). Only one-quarter had had substantial experience judging NDT rounds. Three-quarters of the subjects had little coaching experience, although nearly half (42.1%) had substantial experience debating (5-8 semesters). Two-thirds are university faculty, but for the most part not in the communication-related disciplines (36.8% in speech, drama, journalism or communication).

Material:

Work products and the instrument examined in this study included 1) judging philosophies solicited prior to tournaments, 2) ballots completed during competition at tournaments, and 3) a structured questionnaire administered at tournaments following completion of a majority of the rounds (typically after Round Five).

Twenty subjects completed the questionnaire. Philosophy statements for 16 of these respondents were gathered at one of five tournaments from which ballots were obtained or were taken from the 1988 CEDA national tournament philosophy book. Ballots in sufficient number for analysis (six or more) were available for 17 of the 20 subjects.

The study lost 35 percent of available questionnaires because subjects had completed too few ballots or because no judging philosophy statement was available for the critic. Nearly 70 percent of 551 available ballots were also lost, largely because we had no questionnaire for the critic. Ballots were also lost due to an insufficient number of ballots or because no philosophy statement was available.

In all, 13 critics had sufficient quantities of all three measures (questionnaire, philosophy and ballots) to be included in the pilot study. Twenty of their ballots were blank, therefore unused.

Each of the three measures had a unique development history. Questions for the survey were drawn generally from the researchers' personal experience at various levels of debate, and specifically from concerns noted during the Fall 1988 CEDA season. (Two questions came from Buckley 1983.) The sequence of questions and style of respondent selection options were based upon professional marketing experience and graduate coursework in survey research techniques.

Coding worksheets for content analysis of the philosophy statements and ballots included matrices to capture the proportions of presentational versus substantive and critique versus decision criteria elements of critics' written comments. Such data were important to examination of prospective hypotheses. A set of nine discriminants was selected from results of the questionnaire analysis and included on both worksheets. These variables were characterized by high saliency and moderate variance, thus were considered potentially meaningful. Seven were elements which may influence decisions; two were elements which may influence speaker points. Both worksheets also included Buckley's group of candidate paradigms.

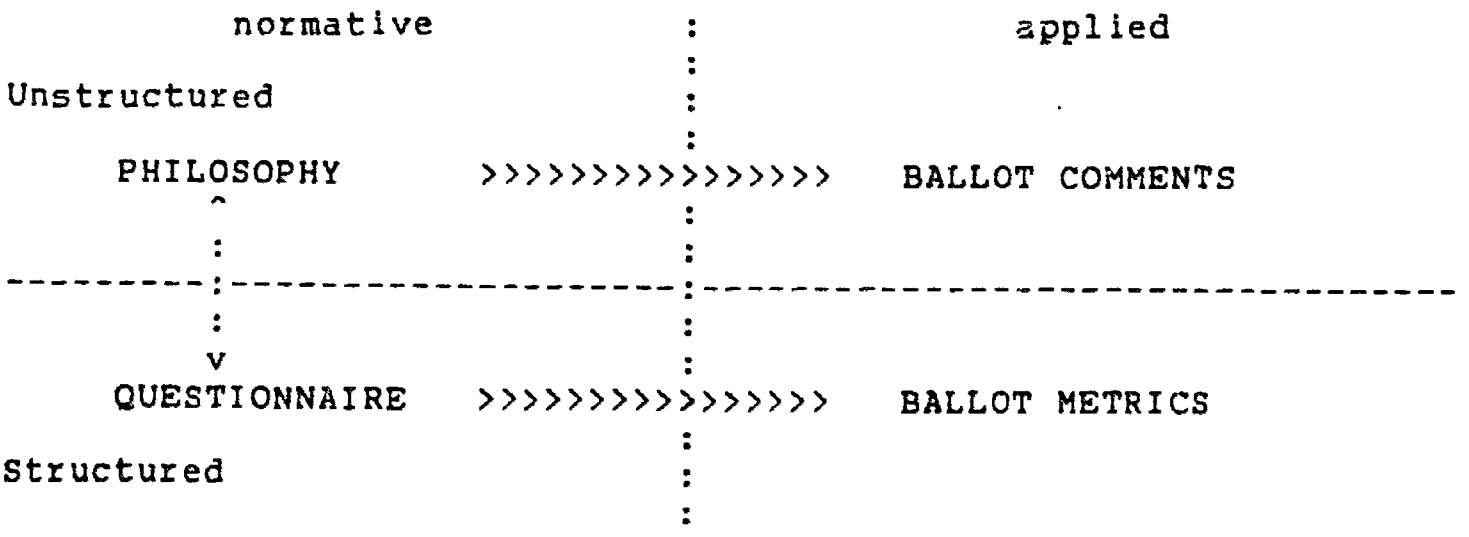
The judging philosophies were assembled first from open-ended essays solicited from critics at the 1989 Syracuse Invitational tournament. This primary format is preferred

because it does not structure or prejudice critics' comments. Although it was necessary to resort to a few philosophies from the CEDA national tournament book, such philosophies are not favored because the large number of items on the CEDA solicitation and its sequence of topics may repress salience. The requirement that CEDA philosophies be completed on a single page may force critics to crowd out their own true priorities in order to make room for answers to specific questions.

The one instrument and two work products used in the study may be visualized conceptually in a two-by-two table. Both the philosophy and questionnaire are normative ("ought") documents; ballots are applied. The philosophy and comment portions of ballots are unstructured; the questionnaire and template (top) portions of ballots are structured. Using these distinctions, future study may examine content, construct and predictive validity of these types of documents.

FIGURE 1

Construct and technique matrix of tools in the study



Procedure:

A two-page questionnaire incorporating 35 Likert Scale items, six yes/no selections, two multiple option questions, and five single-selection choices were administered to judges at CEDA debate tournaments. Two of these questions were repeated from Buckley (1983) in an attempt to replicate partially the earlier study. The questionnaire was administered with little advance publicity, in order to prevent anticipatory modification of ballot decision criteria.

Official ballots submitted by judges at five Spring 1989 CEDA tournaments comprised the second source of data. Each round was considered a unique case for purposes of statistical analysis. One hundred and ninety ballots were analyzed, of which 170 (89.5%) were usable. Ballots analyzed were distributed among the thirteen judges such that each critic's share of the total fell within +/- 50% of random share.

The third source of data were judging philosophy statements, already described. The majority of work products and questionnaires were collected at the Syracuse Debate Invitational held during the last week of January 1989. Additional data were collected at other CEDA tournaments in the Northeast during spring semester (Marist, Richmond, Cornell and William & Mary).

The pool of schools judged in ballots chosen for analysis was influenced somewhat more by rounds involving the U.S. Military Academy (12.9% of affirmatives, 10.0% of negatives) than by debates amongst other schools.

Exactly half the winners were affirmative (and half negative), suggesting that neither the topic nor the critics were biased to one side more than the other.

Nearly two-thirds of the ballots (64.7%) were open division rounds. One-fifth were novice. On the whole, ballots were drawn from midway through their tournaments, ameliorating potential discrepancies between early preliminary round lack of focus by critics and late round fatigue. The sample included a substantial number of elimination rounds (19.9%).

Although AFA Form W dominates the sample (59.4%), the pilot study is too small to support inferences regarding whether the box checkoff style for speaker evaluation may influence judges' criticality or speaker point awards. (In any case, only one critic in 20 made use of the boxes regularly.)

Formal processing began with tabulation and statistical analysis of the questionnaire instrument. A univariate analysis revealed nine discriminants influencing decisions.[1] From this review a set of research concerns was developed. Next, ballot templates were tabulated. Then a content analysis of the judging philosophies and ballot comment sections was conducted. An attempt was made to correlate content variables to elements addressed by the survey. The study also examines the proportion and consistency of comments regarding debaters vs. critiques addressed to the resolution of issues.

All data processing for the pilot study was performed on an IBM PC using PC-FILE PLUS (a database program) and ABC (a statistical package from the University of Michigan). Flexible

data formatting features of the database program were used to grow data definitions as the study progressed, making it possible to add, delete, and rearrange variables as needed without undue loss of time or data. A common fixed-length exchange format allowed swapping of files between the two packages, significantly enhancing the capabilities of both. Generally, data were entered via PC-FILE then exported to ABA for univariate and cross-tabulation runs. The definition of tested database formats and of working procedures for statistical generation is a major benefit of the pilot study that should facilitate later research built upon this effort.

RESULTS

Two research questions and four hypotheses were investigated. The results provide mixed support for the questions and hypotheses.

The first question asked was whether there was a strong relationship between professed reasons for decision and actual reasons cited in debate ballots. Correlational analysis (Pearson's r) found only two instances of professed preference from the questionnaire corresponding with actual preferences.

TABLE 1

Professed vs. actual reasons for decision

	BALLOT COMMENTS					
PROFESSED	CNTRIN	TOPICL	QUALANL	EVCONTX	AFBURDN	EVAPL
Aff fiat key pts	ND	ND	ND	ND	ND	ND

Counterintuitive	ND	.108	ND	ND	ND	ND
Topicality	ND	.028	ND	ND	ND	ND
Qual of Analysis	ND	.050	.152	.698	.001	.166
Ev out of context	ND	.021	.022	.699	.136	.090
Aff Burden of Proof	ND	ND	ND	ND	ND	ND
Applicability of ev	ND	.132	ND	ND	.156	.226

ND = Insufficient Data [2]

The correlations fall below the accepted convention of .80 (Krippendorff 1980) for tentative acceptance. Nevertheless, for a pilot study they provide a limited utility. It appears that for a critic who indicates s/he will vote on "evidence out of context" on a judging philosphy, there is some prospect of finding this actually used on ballots (r = .699). Interestingly, the presence of a statement on "quality of analysis" is also correlated to a similar degree with the presence of "evidence out of context" on ballot comments (r = .698). The low correlations for all other items with sufficient data will be addressed in the Discussion section of the paper.

The second question asked whether there was a strong correlation between judging paradigms and reasons for decision cited in debate ballots. Correlations here indicated that several clusters of ballot behavior were indicated by different paradigms. For instance, there was approximately the same association between judges professing to be Tabula rasa, Value comparison, Argument skills, Hypothesis tester, Judicial model, and Argument critic (range =.698 to .685) and their likelihood

of citing "evidence out of context" in their ballot behavior. Similarly, Value comparison, Argument skills, Judicial model, and Argument critics were about equally likely (range = .674 to .644) to cite "counterintuitive arguments" in their decisions. Finally, those using the Judicial model and Argument critics were similar (range = .589 to .553) in their use of "quality of analysis" in ballot comments. These correlations, while falling below the conventional .80, suggest some judging behaviors which transcend paradigm preference. Other interpretations of the data and their implications are addressed in the Discussion section.

TABLE 2

Paradigms vs. reasons for decision

	BALLOT COMMENTS					
PROFESSED	CNTRIN	TOPICL	QUALANL	EVCONTX	AFBURDN	EVAPL
Tabula Rasa	ND	.043	.049	.698	.316	.098
Value Comparison	.674	.012	.021	.696	.052	.078
Policy Implications	ND	.024	.068	ND	.024	.110
Argument Skills	.669	.012	.010	.694	.127	.095
Argument Critic	.644	.046	.553	.685	.122	.472
Stock Issues	ND	.064	.015	ND	.449	.143
Public Audience	ND	.166	.001	ND	.126	.193
Hypothesis Testing	ND	.089	.025	.693	.207	.156
Judicial Model	.661	.086	.589	.691	.145	.132
Other	ND	.143	.564	ND	.202	.240

ND = Insufficient Data [3]

To determine whether critics displayed consistency in ballot comments, philosophy statements and survey responses, a Critic Consistency Matrix was established. Each discriminant was compared across the two work products and the instrument to determine consistency for the item.[4] The mean consistency rating for the nine items constituted the Critic's Consistency rating. The Critic Consistency Matrix showed moderate to poor levels of consistency for judges across the three measures with a substantial range among critics. The highest consistency value for a judge was a 66.9% consistency rating while the lowest rating was 37.2% (Mean = 54.9).

All three instruments were consistent for only one-seventh (13.7%) of the three measure sets reviewed for each variable. One instrument was consistent with a second (but not the third) measure in 61.5 percent of the cases. Almost one-fourth (24.8%) of the measures were not consistent with either of the other two instruments.

Since the largest proportion of cases found only two of the three instruments in agreement, we evaluated the combination of instruments which were consistent. The results indicate that philosophy statements matched most frequently with another instrument. The low consistency between ballots and surveys implies that the questionnaire responses are a poor predictor of judges' ballot behavior. Philosophy statements, on the other hand, are better predictors of both ballot behavior and survey responses.

TABLE 3

Consistency When Only Two Instruments Matched

MATCHING INSTRUMENTS	% OF TOTAL	UNMATCHED INSTRUMENT
Philosophy, ballot	47.9	Survey
Philosophy, survey	46.5	Ballot
Survey, ballot	5.6	Philosophy

Four research hypotheses were tested. The first hypothesis expected that audience-centered critics would have a higher proportion of presentational (vs. substantive) remarks than analytic-centered critics. Support for this hypothesis was not found. While critics who selected audience-centered paradigms were somewhat more likely to make comments on presentation than critics who choose analytic-centered paradigms, the difference was not significant (P > .05). In fact, for all critics the proportion of comments on presentational elements constituted only about one-sixth of the written comments.

TABLE 4

Hypothesis #1 T-test Values

Presentational Remarks: Audience-centered vs. Analytic-centered

Remark Type	AUD (N=55)		ANA (N=102)		Test	Critical
	Mean	S.D.	Mean	S.D.	Stat	Value
Presentational	1.7	1.7	1.5	1.6	.727	1.960

The second hypothesis predicted audience-centered critics would devote a higher proportion of their ballot to critique (vs. decision criteria) than would analytic-centered critics. This hypothesis also failed to receive support. Audience-centered critics did provide a larger proportion as predicted, but not significantly so (P = > .05). About 43% of audience-centered critics comments were critique compared with about 38% for analytic-centered judges.

TABLE 5

Hypothesis #2 T-test Values

CRITIQUE: Audience-centered vs. Analytic-centered

Remark Type	AUD (N=55)		ANA (N=102)		Test	Critical
	Mean	S.D.	Mean	S.D.	Stat	Value
Critique	3.9	2.6	3.4	2.7	1.115	1.960

The third hypothesis predicted that critics would commit a greater proportion of their elimination round ballots to decision criteria (vs. critique) than they would on their preliminary round ballots. Support was obtained for this hypothesis. Critics devoted almost three-quarters (73.7%) of their elimination round ballots to decision criteria compared with 57.7% of their preliminary round ballots (P = < .05). This result indicates that critics reduce the amount of critique devoted to their written ballots in elimination rounds compared with the amount devoted in preliminary rounds.

TABLE 6

Hypothesis #3 T-test Values

DECISION CRITERIA: Elimination vs. Preliminary rounds						
Remark Type	ELIM (N=34)		PREL (N=136)		Test	Critical
	Mean	S.D.	Mean	S.D.	Stat	Value
Decision Criteria	6.6	2.5	5.2	2.5	2.92	1.960

The fourth hypothesis expected critics to employ more substantive elements (vs. presentational) in their elimination round ballots than in their preliminary round ballots. Support was also obtained for this hypothesis. Over 90% (92.2%) of the elimination round ballots were addresses to substantive issues in the debate. Preliminary round ballots allotted 82.2% of their comments to substantive issues (P = < .05).

TABLE 7

Hypothesis #4 T-test Values

PRESENTATIONAL REMARKS: Elimination vs. Preliminary rounds						
Remark Type	ELIM (N=34)		PREL (N=136)		Test	Critical
	Mean	S.D.	Mean	S.D.	Stat	Value
Substantive	8.3	1.0	7.4	1.5	3.31	1.960

Since the present study drew judging paradigms from Buckley (1983), it was interesting to see the comparative rank-ordering of paradigm preference. Buckley administered questionnaires to

74 critics at four CEDA tournaments. Two of his questions were used in the pilot study questionnaire in an attempt to replicate the earlier research and to suggest a trend line. Analysis of his findings versus our own reveals that in the six years between studies there was little change in rank amongst argument skills, argument critic, hypothesis testing, judicial model, value comparison, and tabula rasa paradigms. Substantial rank changes took place in critics professing policy implications (up), public audience (up), stock issues (down), other (up), and no paradigm (down). However, roughly comparable ranks for the first group may mask substantial proportional changes in argument skills (up), argument critic (down), and value comparison (down). In both studies nearly identical proportions of respondents (approximately 94%) said they would consider criteria from outside their personal paradigm in deciding a round. Due to differences in sample size between the two studies, it is not possible to estimate the validity of apparent trends.

TABLE 8

Professed Paradigms: Buckley vs. Dudczak & Day

	BUCKLEY		DUDCZAK AND DAY	
PARADIGM	Proportion of Critic Responses	Rank	Rank	Proportion of Critic Responses
None	.035	7	11	.000
Argument Skills	.088	5	4	.102
Argument Critic	.123	4	6	.082

Policy Implications	.018	8	2	.125
Hypothesis testing	.053	6	7	.061
Judicial Model	.018	8	10	.041
Game	.000	-	-	not included
Public Audience	.000	11	7	.061
Stock Issues	.211	2	7	.061
Value Comparison	.228	1	2	.125
Tabula Rasa	.211	2	1	.245
Other	.018	8	4	.102

DISCUSSION

The pilot study generated research concerns which will be used to revise the project. An obvious concern is the number of subjects represented in the study. Four potential subjects were lost for their lacking a philosophy statement. Three potential subjects had insufficient ballots. Our intention is to widen the scope of the study from a regional base to a national one. Not only would this increase the number of available subjects, it would also allow cross-regional comparisons. It is otherwise quite possible that regional norms for judging paradigms, philosophy preference, and ballot behavior do not correspond with with national norms. A larger subject pool would generate a substantial increase in ballots (Thirteen subjects produced 170). This in turn should give truer readings for correlations which indicated a direction at sub-significant levels.

We also believe instrumentation adjustments are required.

The questionnaire allowed respondents to select more than one paradigm. Consequently, subjects were often represented (38%) in both the analytic-centered and audience-centered groups. An effect of this would be to minimize differences between group scores. We believe this effect contributed to the failure to support Hypothesis #1 and #2. One modification would be to require subjects to rank and rate paradigm preferences if they selected more than one. This would allow us to construct more discrete groupings.

Another instrument adjustment concerns the variable mix (discriminants). The discriminants on the work products were developed from the questionnaire and selected because they reflected high salience and moderate variability. However, in coding ballots and philosophies, raters identified other variables which were excluded from the work products. Given the poor consistency between the survey instrument and written comments on ballots, generating discriminants from the philosophies and written ballot comments is warranted. Additionally, the exclusion of high variability items may mask bi-modal distributions which correspond with paradigmatic or philosophy preferences. In other words, high standard deviations combined with high salience may indicate divided opinions which cluster in paradigms or philosophies.

A third instrument adjustment would be to develop more standardized definitions for key words in content analysis of the written portion of ballots and philosophy statements. More clearly defined parameters for inclusion (and exclusion) of the

work products would contribute to reliability measures. Of course, the expansion of this pilot would require independent coders and the calculation of an inter-rater reliability quotient.

The researchers were concerned about tournament practices which could skew results. For instance, one tournament advertised an award for the "Best Critic." Another tournament facilitated the collection and distribution of judge philosophy statements. The effect of these and other tournament practices may operate as intervening variables. When such are discovered it seems prudent to conduct a retrospective analysis to determine if an effect could have occurred.

Finally, we believe additional analysis needs to address the predictive and construct validity of work products used for the philosophy statements and written portion of the ballots. These instruments have evolved through the present research project. Their further use in this project and for others requires they establish true relationships between what they purport to measure with that which they actually do measure.

The results of the pilot study are necessarily preliminary. Expanded subject pools will obtain threshold reliability while instrumentation adjustments will accomplish validity. The expanded variable mix should yield more robust differences among critics by their philosophy and paradigm preferences. With the modifications incorporated into the design and instrumentation, the next step for the project will be to begin differentiating profiles among types of judges.

APPENDIX A

Syracuse Debate Union Judging Criteria Questionnaire

Instructions: Please circle a single response for each item.

How much should each of the following influence decisions?

		(never --------> always)				
01.	counter-intuitive arguments	1	2	3	4	5
02.	theoretical arguments	1	2	3	4	5
03.	counter-warrants	1	2	3	4	5
04.	conditional arguments	1	2	3	4	5
05.	evidence attacks	1	2	3	4	5
06.	falsification of evidence	1	2	3	4	5
07.	evidence out of context	1	2	3	4	5
08.	lack of evidence	1	2	3	4	5
09.	acceptability of evidence	1	2	3	4	5
10.	familiarity with evidence	1	2	3	4	5
11.	topicality	1	2	3	4	5
12.	affirmative burden of proof	1	2	3	4	5
13.	quality of analysis	1	2	3	4	5
14.	new arguments in rebuttals	1	2	3	4	5
15.	points made during cross-examination	1	2	3	4	5
16.	adherence to time limits	1	2	3	4	5
17.	affirmative fiat of key case points	1	2	3	4	5
18.	arguments about debating philosophy	1	2	3	4	5
19.	repugnant values	1	2	3	4	5
20.	absence of values	1	2	3	4	5

How should the following elements influence speaker points?

		(never --------> always)				
21.	speed of presentation	1	2	3	4	5
22.	eye contact with judge	1	2	3	4	5
23.	rhetorical pacing	1	2	3	4	5
24.	use of inflection	1	2	3	4	5
25.	gestures	1	2	3	4	5
26.	posture	1	2	3	4	5
27.	obnoxious behavior	1	2	3	4	5
28.	teamwork	1	2	3	4	5

29. Will you discuss decision criteria or speaking style preferences with debaters before a round? Y N

30. Will you discuss your decision or critique debaters immediately after a round? Y N

Which paradigms do you follow, generally? (circle all which apply)

31. Argument Critic
32. Argument Skills
33. Public Audience
34. Hypothesis Testing
35. Tabula Rosa
36. Other (_____)
37. Value Comparison
38. Judicial Model
39. Policy implications
40. Stock Issues
41. one

42. Are you willing to apply criteria from outside your paradigm(s) in rendering a decision? Y N N/A

How important should each judging role be in rendering a decision

		(useless --------> vital)				
43.	issue-related analysis	1	2	3	4	5
44.	decide technical win/loss	1		3	4	5
45.	produce feedback for improvement	1	2	3	4	5

Under what conditions should a judge ask to inspect evidence?
(circle all which apply)

46. never
47. to resolve issues
48. when poor delivery makes evidence unintelligible
49. at judge's option
50. questions of ethics, context
51. when not understood
52. to obtain sources for squad
53. when teams ask

54. Should Affirmative points which are not specifically countered by Negative be held as proven, regardless of inherent strength(s)? Y N

What is the relative importance of these objectives of debate?

		(useless --------> vital)				
55.	Development of speaking skills	1	2	3	4	5
56.	Development of logical reasoning	1	2	3	4	5
57.	Familiarity with research techniques	1	2	3	4	5
58.	Improved crganization	1	2	3	4	5

59. How many years have you judging intercollegiate debate?

0-2 3-5 6-8 9-11 12-14 15-17 18-20 20+

60. What percentage of the rounds you have judged have been NDT?

0-9 10-19 20-29 30-39 40-49 50-59 60-69 70+

61. How many tournament debate rounds have you judged during the past three semesters?

0-16 17-32 33-48 49-64 65-96 97-128 128+

62. How many years have you coached intercollegiate debate?

0-2 3-5 6-8 9-11 12-14 15-17 18-20 20+

63. How many semesters of debating experience have you had personally, in high school and college?

none 1-2 3-4 5-6 7-8 9-10 11-12 13-14 15+

64. Do you hold a degree in either speech, drama, journalism or communications? Y N

65. Do you hold an appointment as a college faculty member (other than as a graduate assistant)? Y N

66. Please print your first and last name. (NOTE: Names will be used for analysis only, not for reporting results.)

First _____ Last _____

APPENDIX B

CODING CATEGORIES FOR BALLOT COMMENTS Acq.# ____

Critic _____ Ballot # _____ Coder _____

I. MATRIX - The written portion of the ballot should be categorized in the following matrix as a percentage of the total in 10% increments:

0 = 0 - 9 %	5 = 50 - 59 %
1 = 10 - 19 %	6 = 60 - 69 %
2 = 20 - 29 %	7 = 70 - 79 %
3 = 30 - 39 %	8 = 80 - 89 %
4 = 40 - 49 %	9 = 90 -100 %

A. Criticism Commentary: Presentation Elements ____

B. Criticism Commentary: Substantive Elements ____

C. Decision Criteria: Presentational Elements ____

D. Decision Criteria: Substantive Elements ____

II. DISCRIMINANTS - Code the following items on the written portion of the ballot:

0 = not present

1 = present in commentary with positive valence

2 = present in commentary with negative valence

3 = present in decision with positive valence

4 = present in decision with negative valence

E. ____ Topicality

F. ____ Quality of Analysis (Analysis)

G. ____ Evidence out of context

H. ____ Affirmative Burden of Proof

I. ____ Applicability of evidence

J. ____ Counter-intuitive arguments

K. ____ Affirmative fiat of key case points

L. ____ Obnoxious behavior (Synonyms: Rude, etc.)

M. ____ Eye Contact with Judge

III. JUDGING PARADIGM - Code each judging paradigm as

1 = mentioned in decision criteria

0 = not mentioned in judging criteria

N. ____ Tabula Rasa

O. ____ Value Comparison

P. ____ Policy Implications

Q. ____ Argument Skills

R. ____ Argument Critic

S. ____ Stock Issues

T. ____ Public Audience

U. ____ Hypothesis Tester

V. ____ Judicial Model

W. ____ Other (_____)

APPENDIX C

JUDGE PHILOSOPHY CODING CATEGORIES Seq.# ____

Critic _____ Coder _____

I. MATRIX - The content of the philosophy should be categorized into two dimensions: Philosophy which deals with "Presentational" elements and that which deals with "Substantive" elements. Use the following range increments:

0 = 0 - 9 %	5 = 50 - 59 %
1 = 10 - 19 %	6 = 60 - 69 %
2 = 20 - 29 %	7 = 70 - 79 %
3 = 30 - 39 %	8 = 80 - 89 %
4 = 40 - 49 %	9 = 90 -100 %

A. Presentational Elements _____

B. Substantive Elements _____

II. DISCRIMINANTS - Code the following items on the philosophy:

0 = not present

1 = mentioned in a positive valence, (i.e., "good," "like," etc.)

2 = mentioned in a negative valence, (i.e., "dislike," "bad," etc.)

C. ____ Topicality

D. ____ Quality of Analysis

E. ____ Evidence out of context

F. ____ Affirmative Burden of Proof

G. ____ Applicability of evidence

H. ____ Counter-intuitive arguments

I. ____ Affirmative fiat of key case points

J. ____ Obnoxious behavior (Synonyms - Rude, etc.

K. ____ Eye Contact with Judge

III. JUDGING PARADIGM - Code each judging paradigm as

0 = not mentioned in philosophy statement

1 = mentioned in philosophy statement

L. ____ Tabula Rasa

M. ____ Value Comparison

N. ____ Policy Implications

O. ____ Argument Skills

P. ____ Argument Critic

Q. ____ Stock Issues

R. ____ Public Audience

S. ____ Hypothesis Tester

T. ____ Judicial Model

U. ____ Other (_____)

IV. SURVEY ITEMS - Code the following items from the Philosphy

0 = not mentioned in philosophy statement

1 = mentioned in a positive valence (i.e., "like," "good," etc.)

2 = mentioned in a negative valence (i.e., "dislike," "bad," etc.)

V. ___ Theoretical Arguments

W. ___ Counter-Warrants

X. ___ Conditional Arguments

Y. ___ Familiarity w/ Evidence

Z. ___ New Args. in Rebuttal

AA. ___ Pts. made in Cross-X

BB. ___ Args. About Debate Phil.

CC. ___ Absence of Values

DD. ___ Speed of Presentation

EE. ___ Teamwork

FF. ___ Discuss Dec. Criteria Before Round

GG. ___ Discuss Decision/Critique After Round

HH. ___ Apply criteria from outside paradigm in decision

II. ___ Accept arguments not countered by opponent (Default)

JJ. ___ Willingness to inspect evidence (after the debate)

ENDNOTES

[1] The nine discriminants were items selected from the questionnaire which had a mean value of > 3.5 or < 2.5 (on a 5 point scale and a standard deviation of 1.0 or less. The nine items were replicated on the coding forms for judging philosophies and ballot content. One item, "Falsification of evidence," met these parameters but was excluded from the discriminant list because its S.D. of 0.00 indicated no variability among subjects. Another item, "Rhetorical pacing, was omitted because anecdotal evidence indicated severe confusion of its definition among respondents to the questionnaire.

[2] Three discriminants were excluded from the correlation. No correlation was found between professed reasons and "affirmative fiat of key case points" as an actual for decision. The other two discriminants, "Obnoxious behavior" and "Eye contact with judge," were presentational elements which were not offered as reasons for decision on the questionnaire.

[3] Six of the 13 respondents included in the pilot study choose more than one paradigm resulting in a 38% overlap between the audience-cen red and analytic-centered subgroups.

[4] Ballot consistency for a subject was calculated as the percentage of ballots indicating the presence of the item. Philosophy statements indicated whether an item was present or not present. Survey responses were rated high (4-5), neutral (3), or low (1-2) for an item.

REFERENCES

Buckley, D. (1983). "A Comparison of Judging Paradigms." ARGUMENT IN TRANSITION: PROCEEDINGS OF THE THIRD SUMMER CONFERENCE ON ARGUMENTATION. Eds. Zarefsky, D., Sillars, M., and Rhodes, J. Annandale, VA: Speech Communication Association, 858-870.

Bryant, M. (1983). "A Content Analysis of Selected NDT and CEDA Transcripts." Unpublished. Paper presented at the Third Biennial Conference on Argumentation. Alta, Utah.

Cox, J. (1974). "A Study of Judging Philosophies of the Participants of the National Debate Tournament." JOURNAL OF THE AMERICAN FORENSIC ASSOCIATION (JAFA) 11: 61-71.

Cross, J., and R. Matlon (1978). "An Analysis of Judging Philosophies in Academic Debate." JAFA 15: 110-123.

Gaske, P., D. Kugler, and J. Theobald (1985). "Judging Attitudes and Paradigmatic Preferences in CEDA Debate: A Cummulative and Construct Validity Investigation." CEDA YEARBOOK 6: 57-66.

Hollihan, T., P. Riley, and C. Austin (1983). "A Content Analysis of Selected CEDA and NDT Judges' Ballots." ARGUMENT IN TRANSITION Eds. Zarefsky, D., Sillars, M., and Rhodes, J. Annandale, VA: Speech Communication Association, 871-882.

Holsti, O. (1969). CONTENT ANALYSIS FOR THE SOCIAL SCIENCES AND HUMANITIES. Reading, MA: Addison-Wesley

Krippendorff, K. (1980). CONTENT ANALYSIS: AN INTRODUCTION TO ITS METHODOLOGY. Beverly Hills: Sage.

Lee, R., K. Lee, and M. Seeger (1983). "Comparison of CEDA and NDT: Differences in Program Characteristics and Forensic Director Attitudes." ARGUMENT IN TRANSITION Eds. Zarefsky, D., Sillars, M., and Rhodes, J. Annandale, VA: Speech Communication Association, 845-857.

Paisley, W. (1969). "Introduction," THE ANALYSIS OF COMMUNICATION CONTENT. Gerbner, G. et al, eds. New York: Wiley, 286.

Thomas, D. (1977) "Analysis of Judges and Debaters at NDT" Unpublished. Data collected from the 1977 National Debate Tournament. Used with permission.

Webb, E. and K. Roberts. (1969) "Unconventional Uses of Content Analysis in Social Science," THE ANALYSIS OF COMMUNICATION CONTENT. Gerbner, G. et al, eds. New York: Wiley, 319-332.